

Vectors and Matrices

Based on Lectures and "Intro to Linear Algebra"

$\theta\omega\theta$

*Not in University of Cambridge
skipped some talks irrelevant to contents*

E-mail: `not telling you`

Contents

I	Complex Numbers	1
1	Definition	1
2	Basic Properties & Consequences	1
3	Exponential and Trigs in \mathbb{C}	2
4	Logarithms and Complex powers	2
5	Transformations, Lines, and Circles	3
II	Vectors in 3 Dimensions	3
6	Vector addition and scalar multiplication	3
7	Dot product	4
8	Vector cross product	4
9	Orthonormal Bases and Components	5
10	Triple products	5
10.1	Scalar triple product	5
10.2	Vector triple product	5
11	Lines, Planes, and Vector equations	6
11.1	Lines	6
11.2	Planes	6
11.3	Other vector equations	6
12	Index notation and the summation convention	7
12.1	Components, δ & ϵ	7
12.2	Summation convention	8
12.3	Applications	8
12.4	ϵ ϵ identity	9
III	Vectors in General	9

13 Vectors in \mathbb{R}^n	9
13.1 Definition and basic properties	9
13.2 Norm of a vector	10
13.3 Cauchy-Schwarz and Triangle inequalities	10
13.4 Inner Products and Cross products	11
14 Vector Spaces	11
14.1 Axioms, span, and subspaces	11
14.2 Linear dependence and independence	12
14.3 Inner products	12
15 Bases and dimension	13

Complex Numbers

1 Definition

Definition 1.1. Construct \mathbb{C} from \mathbb{R} by adding i that $i^2 = -1$. Any $z \in \mathbb{C}$ is in the form

$$z = x + iy, x = \operatorname{Re} z, y = \operatorname{Im} z, x, y \in \mathbb{R}.$$

Addition and multiplication are defined by

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2), z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2).$$

The *conjugate* is defined by

$$\bar{z} = z^* = x - iy.$$

The *modulus* is defined by

$$r = |z|, r \geq 0, r^2 = |z|^2 = z\bar{z} = x^2 + y^2.$$

The *argument* is defined by

$$z \neq 0 : \theta = \arg(z) \in \mathbb{R}, z = r(\cos \theta + i \sin \theta).$$

The values of θ in $(-\pi, \pi]$ are called the *principal values*.

Complex numbers can be plotted on an *Argand diagram*.

2 Basic Properties & Consequences

(1) $+, \times$ are commutative and associative,

\mathbb{C} under $+$ is an abelian group,

\mathbb{C} under \times is an abelian group,

\mathbb{C} is a field.

(2) **Fundamental Theorem of Algebra:** A polynomial with $\deg n$ with coefficients in \mathbb{C} can be written as a product of n linear factors, has at least one solution in \mathbb{C} and has n solutions connected with multiplicity.

(3) Parallelogram constructions.

(4)

$$|z_1| |z_2| = |z_1 z_2|, |z_1 + z_2| \leq |z_1| + |z_2|.$$

Alternative forms:

$$|z_2 - z_1| \geq |z_2| - |z_1|, |z_2 - z_1| \geq ||z_2| - |z_1||.$$

(5) **De Moivre's Theorem:** $z^n = r^n(\cos n\theta + i \sin n\theta)$.

3 Exponential and Trigs in \mathbb{C}

Definition 3.1. Define \exp, \cos, \sin on \mathbb{C} by

$$\begin{aligned}\exp(z) &= e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n, \\ \cos(z) &= \frac{1}{2}(e^{iz} + e^{-iz}) = 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 + \cdots, \\ \sin(z) &= \frac{1}{2i}(e^{iz} - e^{-iz}) = z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 + \cdots.\end{aligned}$$

These series converge for all $z \in \mathbb{C}$. Can be multiplied, rearranged, etc. Definitions reduce to familiar ones in the reals.

Proposition 3.1. $\forall z, w \in \mathbb{C}, e^z e^w = e^{z+w}; e^z e^{-z} = 1, (e^z)^n = e^{nz}, n \in \mathbb{Z}.$

Lemma 3.1. For $z = x + iy$:

- (1) $e^z = e^x(\cos y + i \sin y).$
- (2) $\exp(z) \in \mathbb{C} \setminus \{0\}.$
- (3) $e^z = 1 \Leftrightarrow z = 2\pi ni, n \in \mathbb{Z}.$

Definition 3.2 (Roots of unity). z is an N th root of unity if $z^N = 1$.

We have

$$z^N = r^N e^{iN\theta} = 1 \Leftrightarrow r = 1, N\theta = 2n\pi \Leftrightarrow \theta = \frac{2n\pi}{N},$$

which gives N distinct solutions

$$z = \frac{2n\pi}{N} = \omega^n, \quad n = 0, 1, \dots, N-1.$$

ω^n lie one the vertices of a regular n -gon on the unit circle.

4 Logarithms and Complex powers

Definition 4.1. Define $w = \log z, z \in \mathbb{C} \wedge z \neq 0$ by $e^w = e^{\log z} = z$. Note that since \exp is many-to-one, \log is multi-valued.

$$\begin{aligned}z &= r e^{i\theta} = e^{\log r} e^{i\theta} = e^{\log r + i\theta} \\ \Rightarrow \log z &= \log r + i\theta = \log |z| + i \arg(z)\end{aligned}$$

To make it single-valued, simply take the principal value.

Definition 4.2. Define *complex power* by

$$z^\alpha = e^{\alpha \log z}, \quad z, \alpha \in \mathbb{C}, z \neq 0.$$

Note that since $\arg z \rightarrow \arg z + 2n\pi \Rightarrow z^\alpha \rightarrow z^\alpha e^{2n\pi\alpha}$, it is generally multi-valued. This also reduces to common powers when $z, \alpha \in \mathbb{R}$.

Example.

$$i^i = e^{i \log i} = e^{i(0+i(\frac{\pi}{2}+2n\pi))} = e^{-(\frac{\pi}{2}+2n\pi)}.$$

5 Transformations, Lines, and Circles

- We have five elementary transformations:

$$(1) z \mapsto z + a,$$

$$(2) z \mapsto \lambda z,$$

$$(3) z \mapsto e^{i\alpha} z,$$

$$(4) z \mapsto \bar{z},$$

$$(5) z \mapsto \frac{1}{z}.$$

- General point of a line in \mathbb{C} through z_0 and parallel to w :

$$z = z_0 + \lambda w, \lambda \in \mathbb{R} \text{ or } \bar{w}z - w\bar{z} = \bar{w}z_0 - w\bar{z}_0.$$

- General point of a circle in \mathbb{C} with centre c and radius ρ :

$$z = c + \rho e^{i\theta} \text{ or } |z - c| = \rho \text{ or } |z|^2 - \bar{c}z - c\bar{z} = \rho^2 - |c|^2.$$

- Stereographic projection.

Vectors in 3 Dimensions

6 Vector addition and scalar multiplication

Definition 6.1 (scalar multiplication). Given \mathbf{a} , and scalar $\lambda \in \mathbb{R}$, define $\lambda\mathbf{a}$ to be the position vector of A' on the line OA with length $|\lambda\mathbf{a}| = |\lambda| |\mathbf{a}|$. Direction depends on the sign of λ .

Define $\text{span}\{\mathbf{a}\} = \{\lambda\mathbf{a} : \lambda \in \mathbb{R}\}$. If $\mathbf{a} \neq 0$, then $\text{span}\{\mathbf{a}\}$ is the entire line through O and A .

Define $\mathbf{a} \parallel \mathbf{b}$ if and only if either $\mathbf{a} = \lambda\mathbf{b}$ or $\mathbf{b} = \lambda\mathbf{a}$. Allow $\lambda = 0$, so $\forall \mathbf{a}, \mathbf{0} \parallel \mathbf{a}$. Also allow $\lambda < 0$.

Definition 6.2 (vector addition). Give \mathbf{a}, \mathbf{b} , if $\mathbf{a} \nparallel \mathbf{b}$, construct a parallelogram $OACB$ and define $\mathbf{c} = \mathbf{a} + \mathbf{b}$.

If $\mathbf{a} \parallel \mathbf{b}$, then $\mathbf{a} = \alpha\mathbf{u}, \mathbf{b} = \beta\mathbf{u}$, where \mathbf{u} is a unit vector and $\mathbf{a} + \mathbf{b} = (\alpha + \beta)\mathbf{u}$.

Given $\mathbf{a}, \mathbf{b}, \dots, \mathbf{c}$, we have a linear combination

$$\alpha\mathbf{a} + \beta\mathbf{b} + \dots + \gamma\mathbf{c}$$

for any $\alpha, \beta, \dots, \gamma \in \mathbb{R}$.

Define $\text{span}\{\mathbf{a}, \mathbf{b}, \dots, \mathbf{c}\} = \{\alpha\mathbf{a} + \beta\mathbf{b} + \dots + \gamma\mathbf{c} : \alpha, \beta, \dots, \gamma \in \mathbb{R}\}$. In 3d case, if $\mathbf{a} \nparallel \mathbf{b}$, then $\text{span}\{\mathbf{a}, \mathbf{b}\}$ is a plane through O, A, B .

Here are some properties:

- $\forall \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$, this says that $\mathbf{0}$ is the identity for addition.
- $\exists -\mathbf{a}, \mathbf{a} + (-\mathbf{a}) = (-\mathbf{a}) + \mathbf{a} = \mathbf{0}$. This says $-\mathbf{a}$ is the inverse of \mathbf{a} under addition.
- $\forall \mathbf{a}, \mathbf{b}, \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$, this says that vector addition is commutative.
- $\forall \mathbf{a}, \mathbf{b}, \mathbf{c}, (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$, this says that vector addition is associative.

Hence, the set of vectors with addition form an abelian group.

Relation with scalars:

- $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$.
- $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$.
- $(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a})$.

7 Dot product

Definition 7.1 (dot product). Give \mathbf{a}, \mathbf{b} , let θ be the angle between them, define $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$. Note that θ is defined unless $\mathbf{a} = \mathbf{0}$, in which case we define $\mathbf{a} \cdot \mathbf{b} = 0$. $\mathbf{a} \perp \mathbf{b} \Leftrightarrow \mathbf{a} \cdot \mathbf{b} = 0 \Leftrightarrow \theta = \frac{\pi}{2} \bmod \pi$ when θ is defined. Allow \mathbf{a} or $\mathbf{b} = \mathbf{0}$, so $\mathbf{a} \parallel \mathbf{0} \wedge \mathbf{a} \perp \mathbf{0}$.

For $\mathbf{a} \neq \mathbf{0}$, $|\mathbf{b}| \cos \theta$ is the component of \mathbf{b} along \mathbf{a} .

$$|\mathbf{b}| \cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|} = \mathbf{u} \cdot \mathbf{b}.$$

By resolving \mathbf{b} along and perpendicular to \mathbf{a} , we get

$$\mathbf{b} = \mathbf{b}_{\parallel} + \mathbf{b}_{\perp}.$$

Properties:

- $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}, \mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2 \geq 0, = 0$ iff $\mathbf{a} = \mathbf{0}$.
- $(\lambda\mathbf{a}) \cdot \mathbf{b} = \lambda(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (\lambda\mathbf{b})$.
- $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$.

8 Vector cross product

Definition 8.1. Given \mathbf{a}, \mathbf{b} , let θ be the angle between them, wrt a unit vector \mathbf{n} normal to the plane they span. Define $\mathbf{a} \wedge \mathbf{b}$ or $\mathbf{a} \times \mathbf{b}$ as $|\mathbf{a}||\mathbf{b}| \sin \theta \mathbf{n}$. $\mathbf{0}$ case is similar.

This is the *vector area* of the parallelogram generated by \mathbf{a}, \mathbf{b} . Note that $\mathbf{a} \wedge \mathbf{b} = \mathbf{a} \wedge \mathbf{b}_{\perp}$.

Properties:

- $\mathbf{a} \wedge \mathbf{b} = \mathbf{b} \wedge \mathbf{a}$.
- $(\lambda\mathbf{a}) \wedge \mathbf{b} = \lambda(\mathbf{a} \wedge \mathbf{b}) = \mathbf{a} \wedge (\lambda\mathbf{b})$.
- $\mathbf{a} \wedge (\mathbf{b} + \mathbf{c}) = \mathbf{a} \wedge \mathbf{b} + \mathbf{a} \wedge \mathbf{c}$.
- $\mathbf{a} \wedge \mathbf{b} = \mathbf{0}$ if and only if $\mathbf{a} \parallel \mathbf{b}$.
- $\mathbf{a} \wedge \mathbf{b} \perp \mathbf{a} \wedge \mathbf{b}_{\perp}$.

9 Orthonormal Bases and Components

Choose $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ that are *orthonormal*. That is, they are of unit lengths and $\mathbf{e}_i \cdot \mathbf{e}_j = 0, i \neq j \in \{1, 2, 3\}$, which is equivalent to choose cartesian axes along the directions. Then $\{\mathbf{e}_i\}$ is a basis and $\forall \mathbf{a} \in \mathbb{R}^3$,

$$\mathbf{a} = \sum_{i=1}^3 a_i \mathbf{e}_i.$$

By this spirite, we can write

$$\mathbf{a} = (a_1, a_2, a_3) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

Scalar product in this form can be written as

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3, \quad |\mathbf{a}| = a_1^2 + a_2^2 + a_3^2..$$

For vector products, choose this basis that it is also *right-handed*:

$$\mathbf{e}_i \times \mathbf{e}_{i+1} = \mathbf{e}_{i+2}.$$

Then

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3)(b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3) \\ &= (a_3 b_2 - a_2 b_3) \mathbf{e}_1 + (a_3 b_1 - a_1 b_3) \mathbf{e}_2 + (a_1 b_2 - a_2 b_1) \mathbf{e}_3. \end{aligned}$$

10 Triple products

10.1 Scalar triple product

Definition 10.1. Define scalar triple product by

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}).$$

This is the volumn of the parallelepiped with bases \mathbf{b}, \mathbf{c} and side \mathbf{a} .

Remark. $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ is a "signed" volumn. If $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) > 0$ then $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is called a *right-handed set*. $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ if and only if $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are coplanar, e.g., $\mathbf{c} = \alpha \mathbf{a} + \beta \mathbf{b} \in \text{span}\{\mathbf{a}, \mathbf{b}\}$.

In components,

$$\begin{aligned} \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) &= a_1 b_2 c_3 - a_1 b_3 c_2 \\ &+ a_2 b_3 c_1 - a_2 b_1 c_3 \\ &+ a_3 b_1 c_2 - a_3 b_2 c_1 = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \end{aligned}$$

10.2 Vector triple product

Definition 10.2. Define the vector triple product by $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$. Note that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ does not necessarily give the same result as $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}. \quad (10.1)$$

Lecture 4.

Each component a_i is uniquely determined by $a_i = \mathbf{e}_i \cdot \mathbf{a}$.

We have the following identities:

Proposition 10.1.

$$\begin{aligned}\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) &= \mathbf{0} \\ (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) &= [\mathbf{a}, \mathbf{b}, \mathbf{d}]\mathbf{c} - [\mathbf{a}, \mathbf{b}, \mathbf{c}]\mathbf{d} \\ (\mathbf{a} \times \mathbf{b}) \cdot ((\mathbf{c} \times \mathbf{d}) \times (\mathbf{e} \times \mathbf{f})) &= [\mathbf{a}, \mathbf{b}, \mathbf{d}][\mathbf{c}, \mathbf{e}, \mathbf{f}] - [\mathbf{a}, \mathbf{b}, \mathbf{c}][\mathbf{d}, \mathbf{e}, \mathbf{f}] \\ (\mathbf{b} \times \mathbf{c}) \cdot (\mathbf{a} \times \mathbf{d}) + (\mathbf{c} \times \mathbf{a}) \cdot (\mathbf{b} \times \mathbf{d}) + (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &= 0\end{aligned}$$

11 Lines, Planes, and Vector equations

Vectors are defined as position vectors from O . But the definition of addition enables us to use them to describe displacements between points.

11.1 Lines

General point on a line through \mathbf{a} through \mathbf{u} :

$$\begin{aligned}\mathbf{r} &= \mathbf{a} + \lambda \mathbf{u}, & \lambda \in \mathbb{R} & \text{The parametric form.} \\ \mathbf{u} \times \mathbf{r} &= \mathbf{u} \times \mathbf{a}, & & \text{Cross form.}\end{aligned}$$

Proposition 11.1. Any vector equation of the form $\mathbf{u} \times \mathbf{r} = \mathbf{c}$ represents a line.

Proof. $\mathbf{u} \times \mathbf{r} = \mathbf{c} \Rightarrow \mathbf{u} \cdot (\mathbf{u} \times \mathbf{r}) = \mathbf{u} \cdot \mathbf{c} \Leftrightarrow \mathbf{u} \cdot \mathbf{c} = 0$. If $\mathbf{u} \cdot \mathbf{c} \neq 0$ then the equation is inconsistent. If $\mathbf{u} \cdot \mathbf{c} = 0$, then note that

$$\mathbf{u} \times (\mathbf{u} \times \mathbf{c}) = (\mathbf{u} \cdot \mathbf{c})\mathbf{u} - (\mathbf{u} \cdot \mathbf{u})\mathbf{c} = -|\mathbf{u}|^2 \mathbf{c}.$$

Hence $\mathbf{a} = -(\mathbf{u} \times \mathbf{c})/|\mathbf{u}|^2$ is a solution, and thus it represents a line. \square

11.2 Planes

General point on a plane through \mathbf{a} with directions \mathbf{u}, \mathbf{v} in the plane ($\mathbf{u} \nparallel \mathbf{v}$):

$$\begin{aligned}\mathbf{r} &= \mathbf{a} + \lambda \mathbf{u} + \mu \mathbf{v}, & \lambda, \mu \in \mathbb{R} & \text{Parametric form,} \\ \mathbf{n} \cdot \mathbf{r} &= k = \mathbf{n} \cdot \mathbf{a}, & \mathbf{n} = \mathbf{u} \times \mathbf{v} & \text{Dot form.}\end{aligned}$$

The component of \mathbf{r} along \mathbf{n} is

$$\frac{\mathbf{n} \cdot \mathbf{r}}{|\mathbf{n}|} = \frac{k}{|\mathbf{n}|}.$$

11.3 Other vector equations

$$(1) |\mathbf{r}|^2 + \mathbf{r} \cdot \mathbf{a} = k \Leftrightarrow \left| \mathbf{r} + \frac{1}{2}\mathbf{a} \right|^2 = k + \frac{1}{4}|\mathbf{a}|^2, \text{ a sphere with centre } -\frac{1}{2}\mathbf{a} \text{ and radius } \sqrt{k + \frac{1}{4}|\mathbf{a}|^2}, \text{ provided } k > -\frac{1}{4}|\mathbf{a}|^2.$$

$$(2) \mathbf{r} + \mathbf{a} \times (\mathbf{b} \times \mathbf{r}) = \mathbf{c} \Leftrightarrow \mathbf{r} + (\mathbf{a} \cdot \mathbf{r})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{r} = \mathbf{c}. \text{ Dot with } \mathbf{a}:$$

$$\mathbf{a} \cdot \mathbf{r} = \mathbf{a} \cdot \mathbf{c} \implies (1 - \mathbf{a} \cdot \mathbf{b})\mathbf{r} = \mathbf{c} - (\mathbf{a} \cdot \mathbf{c})\mathbf{b}.$$

If $\mathbf{a} \cdot \mathbf{b} \neq 1$, then there is a unique solution

$$\mathbf{r} = \frac{1}{1 - \mathbf{a} \cdot \mathbf{b}}(\mathbf{c} - (\mathbf{a} \cdot \mathbf{c})\mathbf{b}),$$

which is a point.

If $\mathbf{a} \cdot \mathbf{b} = 1$ and $\text{RHS} \neq 0$, then it is inconsistent.

If $\mathbf{a} \cdot \mathbf{b}$ and $\mathbf{c} - (\mathbf{a} \cdot \mathbf{c})\mathbf{b} = \mathbf{0}$, then

$$(\mathbf{a} \cdot \mathbf{r} - \mathbf{a} \cdot \mathbf{c})\mathbf{b} = \mathbf{0}.$$

Hence it is a plane.

12 Index notation and the summation convention

12.1 Components, δ & ϵ

Write vectors $\mathbf{a}, \mathbf{b}, \dots$ in terms of components a_i, b_i, \dots wrt an orthonormal right-handed basis $\{\mathbf{e}_i\}$. Indices i, j, \dots take values 1, 2, 3.

For example, if $\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}$, then $c_i = [\alpha\mathbf{a} + \beta\mathbf{b}]_i = \alpha a_i + \beta b_i$, for $i = 1, 2, 3$. i is called a *free index*.

Hence

$$\bullet \mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^3 a_i b_i.$$

$$\bullet \mathbf{x} = \mathbf{a} + (\mathbf{b} \cdot \mathbf{c})\mathbf{d} \Leftrightarrow x_j = a_j + \left(\sum_{k=1}^3 b_k c_k \right) d_j.$$

Definition 12.1 (Kronecker delta).

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

We see that $\delta_{ij} = \delta_{ji}$ and also

$$\begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Using this, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$

Definition 12.2 (Levi-Civita epsilon).

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i, j, k) \text{ is an even permutation of } (1, 2, 3), \\ -1 & \text{if } (i, j, k) \text{ is an odd permutation of } (1, 2, 3), \\ 0 & \text{else.} \end{cases}$$

We have $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1, \epsilon_{321} = \epsilon_{213} = \epsilon_{132} = -1$. ϵ_{ijk} is totally anti-symmetric: exchanging any pair of indices produces a change in sign.

Then

$$\mathbf{e}_i \times \mathbf{e}_j = \sum_k \epsilon_{ijk} \mathbf{e}_k$$

and

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= \left(\sum_i a_i \mathbf{e}_i \right) \times \left(\sum_j b_j \mathbf{e}_j \right) \\ &= \sum_{ij} a_i b_j \mathbf{e}_i \times \mathbf{e}_j \\ &= \sum_{ij} a_i b_j \sum_k \epsilon_{ijk} \mathbf{e}_k = \sum_{ijk} a_i b_j \epsilon_{ijk} \mathbf{e}_k\end{aligned}$$

so

$$(\mathbf{a} \times \mathbf{b})_k = \sum_{ij} \epsilon_{ijk} a_i b_j.$$

12.2 Summation convention

With components and index notation, indices that appear twice in a given term are usually summed over. In the summation convention, we omit the sum signs for repeated indices. i.e., the sum is understood.

Example. (i) In $a_i \delta_{ij} = a_1 \delta_{1j} + a_2 \delta_{2j} + a_3 \delta_{3j} = a_j$, since Σ_i is understood.

(ii) $\mathbf{a} \cdot \mathbf{b} = \delta_{ij} a_i b_j = a_i b_i$. Σ_{ij}, Σ_i are understood.

(iii) $(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk} a_j b_k$, Σ_{jk} is understood.

(iv) $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \epsilon_{ijk} a_i b_j c_k$, Σ_{ijk} is understood.

(v) $\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 3$.

(vi)

$$\begin{aligned}[(\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}]_i &= (\mathbf{a} \cdot \mathbf{c})b_i - (\mathbf{a} \cdot \mathbf{b})c_i \\ &= a_j c_j b_i - a_j b_j c_i.\end{aligned}$$

Σ_j is understood.

Here are the rules of summation convention.

- (1) An index occurring exactly once in any given term must appear once in every term in an equation, and it can take any value in 1, 2, 3, a *free* index.
- (2) An index occurring exactly twice in a given term is summed over. A *repeated*, *contracted*, or *dummy* index.
- (3) No index can occur more than twice in any given term.

12.3 Applications

We can use this to prove the vector triple product identity.

Proof. Write the huge sum in summation convention:

$$\begin{aligned}[\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i &= \epsilon_{ijk} a_j (\mathbf{b} \times \mathbf{c})_k \\ &= \epsilon_{ijk} a_j \epsilon_{kpq} b_p c_q \\ &= (\epsilon_{ijk} \epsilon_{kpq}) a_j b_p c_q.\end{aligned}$$

Notice that

$$\epsilon_{ijk} \epsilon_{kpq} = \delta_{ip} \delta_{jq} - \delta_{iq} \delta_{jp} \quad (*)$$

see next subsection. So

$$[\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i = \delta_{ip} \delta_{jq} a_j b_p c_q - \delta_{iq} \delta_{jp} a_j b_p c_q.$$

Notice also that $a_i \delta_{ij} = a_j$, so

$$[\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i = a_q b_i c_q - a_j b_j c_i = (\mathbf{a} \cdot \mathbf{c}) b_i - (\mathbf{a} \cdot \mathbf{b}) c_i.$$

Hence the equation 10.1 is proved. \square

12.4 $\epsilon \epsilon$ identity

Proposition 12.1. $\epsilon_{ijk} \epsilon_{pqk} = \delta_{ip} \delta_{jq} - \delta_{iq} \delta_{jp} = \epsilon_{kij} \epsilon_{kpq}$.

Proof. Notice that LHS and RHS are both anti-symmetric, so both vanish when i, j or p, q take the same value. Inspection shows that¹ it suffices to show the cases $i = p = 1 \wedge j = q = 2$ or $i = q = 1, j = p = 2$ and all other index changings that give non-zero results. \square

¹ Think carefully here.

Proposition 12.2. $\epsilon_{ijk} \epsilon_{pjk} = 2\delta_{ip}$.

Proof. Take $q = j$ in the above equation:

$$\epsilon_{ijk} \epsilon_{pjk} = \delta_{ip} \delta_{jj} - \delta_{ij} \delta_{jp} = 3\delta_{ip} - \delta_{ip} = 2\delta_{ip}.$$

\square

Proposition 12.3. $\epsilon_{ijk} \epsilon_{ijk} = k$.

Proposition 12.4.

$$\begin{aligned} \epsilon_{ijk} \epsilon_{pqr} &= \delta_{ip} \delta_{jq} \delta_{kr} - \delta_{jp} \delta_{iq} \delta_{kr} \\ &\quad + \delta_{jp} \delta_{kq} \delta_{ir} - \delta_{kp} \delta_{jq} \delta_{ir} \\ &\quad + \delta_{kp} \delta_{iq} \delta_{jr} - \delta_{ip} \delta_{kq} \delta_{jr}. \end{aligned}$$

Proof. Total anti-symmetry² in i, j, k and independently in p, q, r implies LHR, RHS agree up to an overall factor. To check the factor is 1, consider $i = p = 1, j = q = 2, k = r = 3$. \square

² This simplifies most of the process and leaves only one case to check.

Vectors in General

13 Vectors in \mathbb{R}^n

13.1 Definition and basic properties

Lecture 6

Definition 13.1. Regard vectors as sets of components, and let

$$\mathbb{R}^n = \{\mathbf{x} = (x_1, \dots, x_n) : x_i \in \mathbb{R}\}.$$

Define:

- Addition: $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n)$,
- Scalar multiplication: $\lambda \mathbf{x} = (\lambda x_1, \dots, \lambda x_n)$.
- Linear combinations: $\lambda \mathbf{x} + \mu \mathbf{y}$,
- Parallel: $\mathbf{x} \parallel \mathbf{y} \Leftrightarrow \mathbf{x} = \lambda \mathbf{y} \vee \mathbf{y} = \lambda \mathbf{x}$.
- Inner Product (Scalar product): $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i$.

Properties of inner product:

(1). Symmetric: $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.

(2). Bilinear:

$$\begin{aligned} (\lambda \mathbf{x} + \lambda' \mathbf{x}') \cdot \mathbf{y} &= \lambda \mathbf{x} \cdot \mathbf{y} + \lambda' \mathbf{x}' \cdot \mathbf{y}, \\ \mathbf{x} \cdot (\mu \mathbf{y} + \mu' \mathbf{y}') &= \mu \mathbf{x} \cdot \mathbf{y} + \mu' \mathbf{x} \cdot \mathbf{y}'. \end{aligned}$$

(3). Positive definite: $\mathbf{x} \cdot \mathbf{x} \geq 0$, with $=$ holds if and only if $\mathbf{x} = \mathbf{0}$.

13.2 Norm of a vector

Definition 13.2. The *norm* of a vector \mathbf{x} is denoted as $|\mathbf{x}|$ with $|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x}$. \mathbf{x}, \mathbf{y} are called *orthogonal* if $\mathbf{x} \cdot \mathbf{y} = 0$, denote as $\mathbf{x} \perp \mathbf{y}$.

The *standard basis* of \mathbb{R}^n is

$$\mathbf{e}_i = (0, \dots, 1, \dots, 0)$$

with 1 on the i th position. So that

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$$

and $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. i.e., standard basis is orthogonal.

13.3 Cauchy-Schwarz and Triangle inequalities

Proposition 13.1 (Cauchy-Schwarz).

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad |\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}| |\mathbf{y}|$$

with equality if and only if $\mathbf{x} \parallel \mathbf{y}$.

General deductions:

(i). Setting $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}| |\mathbf{y}| \cos \theta$, we can define angle θ between $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(ii). We have the *triangle inequality*:

$$|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|.$$

Proof. If $\mathbf{y} = \mathbf{0}$, then the result is immediate. If not, consider

$$\begin{aligned} |\mathbf{x} - \lambda\mathbf{y}|^2 &= (\mathbf{x} - \lambda\mathbf{y}) \cdot (\mathbf{x} - \lambda\mathbf{y}) \\ &= |\mathbf{x}|^2 - 2\lambda\mathbf{x} \cdot \mathbf{y} + \lambda^2|\mathbf{y}|^2 \geq 0. \end{aligned}$$

This is a real equation of λ with at most one root, so

$$(-2\mathbf{x} \cdot \mathbf{y})^2 - 4|\mathbf{x}|^2|\mathbf{y}|^2 \leq 0 \iff |\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|.$$

Equality holds if and only if $\mathbf{x} = \lambda\mathbf{y}$.

Note also that for triangle inequality:

$$\begin{aligned} |\mathbf{x} + \mathbf{y}|^2 &= |\mathbf{x}|^2 + 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2 \\ &\leq |\mathbf{x}|^2 + 2|\mathbf{x}||\mathbf{y}| + |\mathbf{y}|^2 \\ &= (|\mathbf{x}| + |\mathbf{y}|)^2, \end{aligned}$$

as required. □

13.4 Inner Products and Cross products

Inner product in \mathbb{R}^n can be written as

$$\mathbf{a} \cdot \mathbf{b} = \delta_{ij}a_i b_j, \quad \text{by summation convention.}$$

For $n = 3$, it matches geometrical definition.

We can also define cross product in component definition. In 3d we have

$$(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk}a_j b_k,$$

and in n dimensions we have $\epsilon_{ij\dots l}$ which is totally anti-symmetric. But there are only two $a_i b_j$ so we cannot use this to define vector product in general.

However, in \mathbb{R}^2 we have ϵ_{ij} with $\epsilon_{12} = -\epsilon_{21} = 1$, so can use this to define a new scalar product

$$[\mathbf{a}, \mathbf{b}] = \epsilon_{ij}a_i b_j = a_1 b_2 - a_2 b_1.$$

Geometrically, this is the (signed) area of parallelogram formed by \mathbf{a}, \mathbf{b} and

$$|[\mathbf{a}, \mathbf{b}]| = |\mathbf{a}||\mathbf{b}|\sin\theta.$$

Compare with $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \epsilon_{ijk}a_i b_j c_k$.

14 Vector Spaces

14.1 Axioms, span, and subspaces

Definition 14.1. Let V be a set of objects called *vectors* with operation

$$\begin{aligned} \mathbf{v} + \mathbf{w} &\in V \quad \forall \mathbf{v}, \mathbf{w} \in V \\ \lambda \mathbf{v} &\in V \quad \forall \mathbf{v} \in V, \lambda \in \mathbb{R}. \end{aligned}$$

Then V is called a *real vector space* if

- (i). V with $+$ is an abelian group.

$$(ii). \lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$$

$$(iii). (\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$$

$$(iv). \lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$$

$$(v). 1\mathbf{v} = \mathbf{v}.$$

Example. Let $V = \{f : [0, 1] \rightarrow \mathbb{R} : f \text{ is smooth and } f(0) = f(1) = 0\}$. By smooth we mean f is differentiable infinitely many times. Then V is a real vector space with $+$ defined as $(f + g)(x) = f(x) + g(x)$ and $(\lambda f)(x) = \lambda(f(x))$. Then all axioms apply.

Definition 14.2. A *subspace* of a real vector space V is a subset $U \subseteq V$ that is also a vector space.

Lecture 7

Remark. A non-empty subset is a subspace if and only if $\forall v, w \in U, \lambda v + \mu w \in U$.

For any vectors $v_1, \dots, v_r \in V$, their *span* $\text{span}\{v_1, \dots, v_r\} = \{\lambda_1 v_1 + \dots + \lambda_r v_r : \lambda_i \in \mathbb{R}\}$ is a subspace. V and $\{0\}$ are subspaces of V .

Example. A line or plane through O is a subspace in \mathbb{R}^3 , but a line or plane that does not contain O is not a subspace.

14.2 Linear dependence and independence

For $v_1, \dots, v_r \in V$, a real vector space, consider a linear relation

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0. \quad (*)$$

If $(*) \Rightarrow \lambda_i = 0$, then the vectors form a *linearly independent set*. They obey only the trivial linear relation.

If $(*)$ holds with at least $\lambda_k \neq 0$, then the vectors form a *linearly dependent set*. They obey a non-trivial linear relation.

Example. In \mathbb{R}^2 , $\{(1, 0), (0, 1), (0, 2)\}$ is linearly dependent.

We cannot express $(1, 0)$ in terms of the others.

Several facts:

- Any set containing 0 is linearly dependent.
- In \mathbb{R}^3 , $\{\mathbf{a}\}$ is linearly independent if and only if $\mathbf{a} \neq \mathbf{0}$.
- $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is linearly independent if $[\mathbf{a}, \mathbf{b}, \mathbf{c}] \neq 0$. Since if

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = 0,$$

then dotting with $\mathbf{b} \times \mathbf{c}$ we get $\alpha[\mathbf{a}, \mathbf{b}, \mathbf{c}] = 0 \Rightarrow \alpha = 0$. Similarly $\beta = 0, \gamma = 0$.

14.3 Inner products

This is an additional structure on a real vector space V , that can also be characterised by axioms or key properties.

For $v, w \in V$, denote inner product by

$$v \cdot w \text{ or } (v, w) \in \mathbb{R}.$$

Require this satisfies 1. it's symmetric, 2. it's bilinear, 3. it is positive definite.

Definition of length or norm and deductions such as Cauchy-Schwarz inequality depend just on these properties.

Example. Consider space of functions

$$V = \{f : [0, 1] \rightarrow \mathbb{R} : f \text{ smooth} \wedge f(0) = f(1) = 0\}.$$

Define an inner product by

$$(f, g) = \int_0^1 f(x)g(x) \, dx.$$

which has properties 123. Cauchy-Schwarz holds:

$$|(f, g)| \leq \|f\| \|g\|$$

with $\|f\|^2 = (f, f)$. i.e.

$$\left| \int_0^1 f(x)g(x) \, dx \right| \leq \left(\int_0^1 f(x) \, dx \right)^{1/2} \left(\int_0^1 g(x) \, dx \right)^{1/2}.$$

Lemma 14.1. In any real vector space V with an inner product, if v_1, v_2, \dots, v_r are non-zero and orthogonal vectors, then they are linearly independent.

Proof. If

$$\sum_i \alpha_i v_i = 0,$$

then

$$(v_j, \sum_i \alpha_i v_i) = 0 \iff \alpha_j = 0.$$

□

15 Bases and dimension

Definition 15.1. For a vector space V , a *basis* is a set

$$\mathfrak{B} = \{e_1, \dots, e_n\}$$

such that

(i) \mathfrak{B} spans V . i.e., $\forall v \in V$,

$$v = \sum_{i=1}^n v_i e_i.$$

(ii) \mathfrak{B} is linearly independent.

Given (ii), the coefficients v_i in (i) are unique, since

$$\sum_i v_i e_i = \sum_i v'_i e_i \iff v_i - v'_i = 0 \iff v_i = v'_i.$$

Example. Standard basis for \mathbb{R}^n consists of

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Many other bases can be chosen.

Theorem 15.1. If $\{e_1, \dots, e_n\}, \{f_1, \dots, f_m\}$ are bases for a real vector space V , then $m = n$.

Proof. We have

$$\begin{aligned} f_a &= \sum_i A_{ai} e_i, \\ e_i &= \sum_a B_{ia} f_a \end{aligned}$$

for $A_{ai}, B_{ia} \in \mathbb{R}$. Hence

$$\begin{aligned} f_a &= \sum_i A_{ai} \sum_b B_{ib} f_b \\ &= \sum_b \sum_i A_{ai} B_{ib} f_b \end{aligned}$$

But the coefficients are unique, so

$$\sum_i A_{ai} B_{ib} = \delta_{ab}.$$

Similarly,

$$\sum_a B_{ia} A_{aj} = \delta_{ij}.$$

Now,

$$\sum_{i,a} A_{ai} B_{ia} = \sum_a \delta_{aa} = m = \sum_i \delta_{ii} = n.$$

□

Definition 15.2. The number of vectors in any basis is the *dimension* of the vector space.